

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

JP05300213
VOICE DIALER
SHARP CORP

Inventor(s): KAMIYA SHIN

Application No. 04100930 , Filed 19920421 , Published 19931112

Abstract:

PURPOSE: To obtain the voice dialer able to retrieve an object word standard pattern in a short time even when a registered word is forgotten without need for designation of a user number.

CONSTITUTION: A word recognition section 1 obtains a shortest matching distance WD_k of a characteristic pattern of an input voice signal from a microphone 11. A talker recognition section 8 obtains a shortest matching distance SP_k of a speaker characteristic pattern of the inputted voice signal similarly. When the shortest matching distance WD_k is a threshold level TWD or over, the control section 7 confirms it that a word standard pattern relating to a speaker standard pattern representing the shortest matching distance SP_k is an object word standard pattern. A voice analysis/synthesis section 3 generates an echo back sound of the confirmed word standard pattern. Thus, even when a name of a registered word is forgotten, the object word standard pattern is retrieved in a short time by having only to make any utterance without designation of the user number so as to limit the retrieval scope to the word standard pattern of the uttered user.

COPYRIGHT: (C)1993,JPO&Japio

Int'l Class: H04M00127 G10L00300 G10L00300

MicroPatent Reference Number: 001937212

COPYRIGHT: (C) JPO

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平5-300213

(43)公開日 平成5年(1993)11月12日

(51)Int.Cl.⁵

H 0 4 M 1/27

G 1 0 L 3/00

識別記号

5 3 1 Z

5 5 1 A

庁内整理番号

7190-5K

8842-5H

8842-5H

F I

技術表示箇所

審査請求 未請求 請求項の数 2 (全 10 頁)

(21)出願番号

特願平4-100930

(22)出願日

平成4年(1992)4月21日

(71)出願人 000005049

シャープ株式会社

大阪府大阪市阿倍野区長池町22番22号

(72)発明者 神谷 伸

大阪府大阪市阿倍野区長池町22番22号

シャープ株式会社内

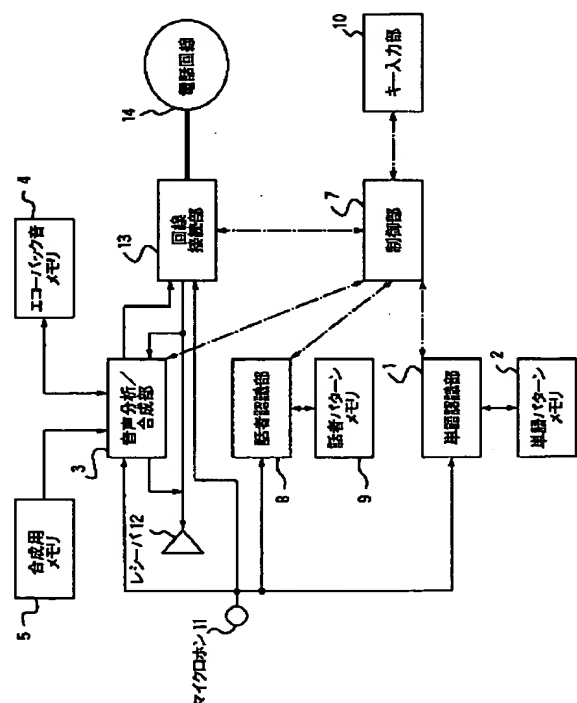
(74)代理人 弁理士 青山 葆 (外1名)

(54)【発明の名称】 音声ダイヤラ

(57)【要約】

【目的】 ユーザ番号指定の必要がなく、且つ登録単語を忘れても目的とする単語標準パターンを短時間に検索できる音声ダイヤラを提供する。

【構成】 単語認識部1は、マイクロホン11からの入力音声の特徴パターンの最短マッチング距離 WD_k を得る。話者認識部8は、同じ入力音声の話者特徴パターンの最短マッチング距離 SP_k を得る。制御部7は、最短マッチング距離 WD_k が閾値 TWD 以上である場合には、最短マッチング距離 SP_k を呈する話者標準パターンに関連付けられた単語標準パターンを目的とする単語標準パターンであると確定する。音声分析/合成部3は、確定された単語標準パターンのエコーバック音を生成する。こうして、登録単語を忘れた場合でも、ユーザ番号を指定することなく、何らかの発声を行えば検索範囲を発声ユーザの単語標準パターンに絞って目的とする単語標準パターンを短時間に検索できる。



【特許請求の範囲】

【請求項 1】 マイクロホンに入力された音声を認識して認識結果に対応した電話番号を表す信号を出力する音声ダイヤラであって、

電話番号に対応付けて登録する登録単語の音声信号に基づく単語標準パターンを格納する単語パターンメモリと、

上記マイクロホンから入力された音声信号を分析して得られた特徴パターンと上記単語パターンメモリに格納された各単語標準パターンとのマッチング距離を算出し、少なくとも最小マッチング距離を呈する単語標準パターンを入力音声の認識結果とする音声認識部と、

上記電話番号に対応付けて登録する登録単語の音声信号に基づく話者標準パターンを、上記単語標準パターンに関連付けて格納する話者パターンメモリと、

上記マイクロホンから入力された音声信号を分析して得られた話者特徴パターンと上記話者パターンメモリに格納された各話者標準パターンとのマッチング距離を算出し、少なくとも最小マッチング距離を呈する話者標準パターンを入力音声の話者の認識結果とする話者認識部と、

上記音声認識部による認識結果に係る最小マッチング距離が所定値より小さい場合には、その認識結果を入力音声の正しい認識結果であると確定する一方、上記音声認識部による認識結果に係る最小マッチング距離が上記所定値以上である場合には、上記話者認識部による認識結果に関連付けられた単語標準パターンを上記入力音声の正しい認識結果であると確定する認識結果確定部と、上記認識結果確定部によって確定された単語標準パターンの登録単語を出力する認識結果出力部を備えたことを特徴とする音声ダイヤラ。

【請求項 2】 請求項 1 に記載の音声ダイヤラにおいて、

上記音声認識部は、上記マッチング距離の小さい順に複数の単語標準パターンを認識結果とし、

上記話者認識部は、上記マッチング距離の小さい順に複数の話者標準パターンを認識結果とし、

上記認識結果出力部は、上記認識結果確定部によって入力音声の正しい認識結果であると確定された複数の単語標準パターンの登録単語を認識候補として出力することを特徴とする音声ダイヤラ。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は、音声認識機能および音声分析/合成機能を利用した音声ダイヤラに関する。

【0002】

【従来の技術】 近年、音声認識機能を利用した音声ダイヤラを備えた電話機が種々提案されている。上記音声ダイヤラは、予め電話番号に対応付けて登録する登録単語を発声し、音声認識部によって入力音声の特徴パターン

を抽出して単語標準パターンとして電話番号に対応付けてメモリに登録しておく。そして、発信時には、目的とする電話番号に対応付けられている登録単語を発声し、上記音声認識部での単語標準パターンとのマッチングによる認識結果に対応付けられた電話番号を電話回線に発信するようにしている。

【0003】ところで、上述のように任意の電話番号に対応付けられてメモリに登録された登録単語の単語標準パターンを消去したり修正したりする場合には、入力部の所定キーを操作して選択した単語標準パターンが消去あるいは修正したい単語標準パターンであるか否かを確認する必要がある。そこで、上記選出された単語標準パターンにリンクされたエコーバック音用パラメータを読み出して、エコーバック音に再生して出力するようにしている。

【0004】また、消去あるいは修正したい単語標準パターンの他の選出方法として、目的とする単語標準パターンに係る登録単語を発声し、音声認識部によって入力音声を認識することによって消去/修正の対象となる音声標準パターンを指定する方法が提案されている(特開平 2-193195 号公報)。

【0005】

【発明が解決しようとする課題】 しかしながら、上記従来の単語標準パターンの消去/修正方法には、1 台の音声ダイヤラ付電話機を複数のユーザが用いる場合には以下のような問題点が生ずる。すなわち、上述のエコーバック音に従って所望の単語標準パターンを検索する消去/修正方法においては、例えば単語標準パターンをユーザ番号に対応付けて上記メモリに登録することによって単語標準パターンを各ユーザ毎に管理している場合には、ユーザ番号をキー入力等によって指定して予め検索対象の単語標準パターンを絞ることが可能であるから検索時間上さほど問題はない。ところが、ユーザ番号指定操作を省くために複数ユーザの単語標準パターンを同じ領域に混在させて登録しておく場合には、先頭の単語標準パターンから順次エコーバック音によって検索しなければならないので検索に時間が掛かるという問題がある。

【0006】また、所望の単語標準パターンに係る登録単語を発声して音声認識部による認識処理によって所望の単語標準パターンを検索する消去/修正方法においては、ユーザが消去/修正したい単語標準パターンに係る登録単語を忘れた場合には、当て推量によって発声を繰り返して又は上述のようにエコーバック音によって検索しなければならず、いずれにしても検索に時間が掛かるという問題がある。

【0007】そこで、この発明の目的は、ユーザ番号指定の必要がなく、且つ登録単語を忘れても音声入力によって目的とする単語標準パターンを短時間に検索できる音声ダイヤラを提供することにある。

【0008】

【課題を解決するための手段】上記目的を達成するため、第1の発明の音声ダイヤラは、マイクロホンに入力された音声認識して認識結果に対応した電話番号を表す信号を出力する音声ダイヤラであって、電話番号に対応付けて登録する登録単語の音声信号に基づく単語標準パターンを格納する単語パターンメモリと、上記マイクロホンから入力された音声信号を分析して得られた特徴パターンと上記単語パターンメモリに格納された各単語標準パターンとのマッチング距離を算出して少なくとも最小マッチング距離を呈する単語標準パターンを入力音声の認識結果とする音声認識部と、上記電話番号に対応付けて登録する登録単語の音声信号に基づく話者標準パターンを上記単語標準パターンに関連付けて格納する話者パターンメモリと、上記マイクロホンから入力された音声信号を分析して得られた話者特徴パターンと上記話者パターンメモリに格納された各話者標準パターンとのマッチング距離を算出して少なくとも最小マッチング距離を呈する話者標準パターンを入力音声の話者の認識結果とする話者認識部と、上記音声認識部による認識結果に係る最小マッチング距離が所定値より小さい場合にはその認識結果を入力音声の正しい認識結果であると確定する一方、上記音声認識部による認識結果に係る最小マッチング距離が上記所定値以上である場合には上記話者認識部による認識結果に関連付けられた単語標準パターンを上記入力音声の正しい認識結果であると確定する認識結果確定部と、上記認識結果確定部によって確定された単語標準パターンの登録単語を出力する認識結果出力部を備えたことを特徴としている。

【0009】また、第2の発明の音声ダイヤラは、第1の発明の音声ダイヤラにおいて、上記音声認識部は上記マッチング距離の小さい順に複数の単語標準パターンを認識結果とし、上記話者認識部は上記マッチング距離の小さい順に複数の話者標準パターンを認識結果とし、上記認識結果出力部は上記認識結果確定部によって入力音声の正しい認識結果であると確定された複数の単語標準パターンの登録単語を認識候補として出力することを特徴としている。

【0010】

【作用】第1の発明では、マイクロホンに向かって電話番号に対応付けられて登録されている登録単語等の単語が発声されると、上記マイクロホンからの音声信号が単語認識部および話者認識部に入力される。そして、上記単語認識部によって、上記音声信号を分析して得られた特徴パターンと単語パターンメモリに格納された各単語標準パターンとのマッチング距離が算出され、少なくとも最小マッチング距離を呈する単語標準パターンが認識結果として出力される。一方、上記話者認識部によって、上記音声信号を分析して得られた話者特徴パターンと話者パターンメモリに格納された各話者標準パターン

とのマッチング距離が算出され、少なくとも最小マッチング距離を呈する話者標準パターンが認識結果として出力される。

【0011】そうすると、認識結果確定部によって、上記音声認識部による認識結果に係る最小マッチング距離が所定値より小さい場合には、その認識結果が入力音声の正しい認識結果であると確定される。一方、上記音声認識部による認識結果に係る最小マッチング距離が上記所定値以上である場合には、上記話者認識部による認識結果に関連付けられた単語標準パターンが上記入力音声の正しい認識結果であると確定される。そして、上記認識結果確定部によって正しい認識結果が確定されると、この確定された単語標準パターンの登録単語が認識結果出力部によって出力される。

【0012】したがって、目的とする登録単語を忘れたユーザが上記マイクロホンに向かって上記目的とする登録単語の発音に近い発音の単語を発声した場合には、上記話者認識部によって上記ユーザの発声による話者標準パターンのうち上記発声単語に最も近い発音の登録単語（すなわち、目的とする登録単語）の話者標準パターンが選出される。こうして、上記目的とする登録単語を忘れた場合であっても、自動的に上記目的とする登録単語の単語標準パターンの登録単語が上記認識結果出力部によって出力されるのである。

【0013】また、第2の発明では、上記音声認識部によって上記マッチング距離の小さい順に複数の単語標準パターンが認識結果として出力される。一方、上記話者認識部によって上記マッチング距離の小さい順に複数の単語標準パターンが認識結果として出力される。そうすると、上記認識結果確定部によって、上記音声認識部による複数の認識結果および上記話者認識部による複数の認識結果に基づいて、複数の単語標準パターンが上記入力音声の正しい認識結果であると確定される。そして、この確定された複数の単語標準パターンの登録単語が認識候補として認識結果出力部によって出力される。

【0014】したがって、目的とする登録単語を忘れたユーザが上記マイクロホンに向かって出た目な単語を発声した場合には、上記話者認識部によって上記ユーザの発声による上記目的とする登録単語を含む複数の話者標準パターンが選出される。こうして、上記目的とする登録単語を忘れた場合であっても、自動的に上記目的とする登録単語を含む複数の認識候補が上記認識結果出力部によって出力されるのである。

【0015】

【実施例】以下、この発明を図示の実施例により詳細に説明する。図1は、本実施例の音声ダイヤラのブロック図である。この音声ダイヤラは、上記音声認識部としての単語認識部1、単語パターンメモリ2、上記認識結果出力部としての音声分析/合成部3、エコーバック音メモリ4、話者認識部8、話者パターンメモリ9、回線接続部1

3および制御部7等から概略構成される。

【0016】上記単語認識部1は、電話番号に対応付けて登録する登録単語の音声マイクロホン11から入力されると、この入力された音声信号を分析して単語単位の特徴パラメータの時系列を得る。そして、この特徴パラメータの時系列を単語標準パターンとして単語パターンメモリ2に記憶する。また、上記マイクロホン11から入力された音声信号を分析して得られた上記特徴パラメータ時系列から成る特徴パターンと、上述のようにして予め単語パターンメモリ2に格納されている複数の単語標準パターンとのマッチングを行う。そして、最も近いマッチング距離を有する単語標準パターンの番号kおよび両パターン間の距離 WD_k を制御部7に送出する。

【0017】上記音声分析/合成部3は、マイクロホン11から入力された音声信号の波形を分析して圧縮し、得られたパラメータをエコーバック音用パラメータとしてエコーバック音メモリ4に格納する。また、上記音声分析/合成部3は、エコーバック音メモリ4に格納されたエコーバック音用パラメータあるいは合成用メモリ5に格納された音声合成用パラメータを読み出して音声信号を生成する。こうして生成された合成音声による音声ガイダンスあるいはエコーバック音はレシーバ12から出力されるのである。

【0018】上記話者認識部8は、上記登録単語の音声上記マイクロホン11から入力されると、この入力された音声信号を分析して単語単位の特徴パラメータ(平均ピッチ周波数、平均パワースペクトルの傾き等)の時系列を得る。そして、この特徴パラメータ時系列を話者標準パターンとして話者パターンメモリ9に記憶する。また、上記マイクロホン11から入力された音声信号を分析して得られた上記特徴パラメータ時系列から成る話者特徴パターンと、上述のようにして予め話者パターンメモリ9に格納されている複数の話者標準パターンとのマッチングを行う。そして、最も近いマッチング距離を有する話者標準パターンの番号kおよび両パターン間の距離 SP_k を制御部7に送出する。

【0019】上記回線接続部13は、送信時には、上記制御部7より相手の電話番号を表す信号が入力されるとその電話番号を電話回線14に送信する。こうして、相手の電話機と電話回線が接続された後は、マイクロホン11から入力された音声信号を電話回線14に送信するのである。また、受信時には、上記電話回線14から受信した音声信号をレシーバ12に送出する。

【0020】上記制御部7は、上述の単語認識部1、音声分析/合成部3、話者認識部8および回線接続部13等を制御して、音声入力によって指定された電話番号を電話回線14に発信したり、消去/修正の対象となる単語標準パターンを検索したりする。また、キー入力部10からは、上記制御部7に対する指示等が入力される。

【0021】上記構成の音声ダイヤラは次のように動作

する。本音声ダイヤラの動作モードには、予め上記単語標準パターンや話者標準パターンを登録する登録モードと入力された音声認識する認識モードとの2つのモードがある。この2つの動作モードおよび後に詳述する他の動作モードの切り替え制御は、制御7によって行われる。以下、各動作モードについて詳細に説明する。

【0022】＜登録モード＞この登録モードは、さらに単語標準パターン登録モードおよび話者標準パターン登録モードの2つの動作モードに分かれる。

〔単語標準パターン登録モード〕ユーザは、合成用メモリ5に格納された音声合成用パラメータに基づいて音声分析/合成部3によって生成された合成音声による音声ガイダンスに従って、電話番号に対応付けて登録する登録単語をマイクロホン11に向かって発声する。そして、上記マイクロホン11から出力された音声信号は単語認識部1、音声分析/合成部3および話者認識部8に入力され、話者認識部8に入力された音声信号はA/D変換された後バッファに一旦格納される。

【0023】そうすると、上記単語認識部1においては、上述のようにして得られた単語標準パターンを単語パターンメモリ2に登録する。同時に、上記音声分析/合成部3においては、上述のようにして得られたエコーバック音用パラメータをエコーバック音メモリ4に格納する。その際に、上記制御部7は、上記単語パターンメモリ2に格納された単語標準パターンとエコーバック音メモリ4に格納されたエコーバック音用パラメータとのリンク情報を内部メモリ(図示せず)に格納する。また、上記単語標準パターンとキー入力部10からキー入力された電話番号とのリンク情報をも上記内部メモリに格納する。

【0024】こうして、上記単語標準パターンの登録が終了すると、制御部7によって動作モードが上記話者標準パターン登録モードに切り替えられる。

【0025】〔話者標準パターン登録モード〕そうすると、上記話者認識部8においては、上述のようにバッファに格納されている登録単語のデジタル音声信号に基づいて上述のようにして得られた話者標準パターンを話者パターンメモリ9に登録する。その際に、上記制御部7は、上記話者パターンメモリ9に格納された話者標準パターンと単語パターンメモリ2に格納された単語標準パターンとのリンク情報を上記内部メモリに格納する。

【0026】このようにして、上記単語標準パターンおよび話者標準パターンの登録が終了すると、制御部7によって動作モードが上記認識モードに切り替えられる。＜認識モード＞この認識モードは、さらに単語認識モードおよび話者認識モードの2つの動作モードに分かれる。

【0027】〔単語認識モード〕ユーザは、上記マイクロホン11に向かって、所望する電話番号に対応付けられている登録単語または消去/修正の対象となる登録単

語を発声する。そうすると、マイクロホン11からの音声信号は単語認識部1および話者認識部8に入力され、話者認識部8に入力された音声信号はA/D変換された後上記バッファに一旦格納される。そして、単語認識部1においては上記マッチングを行い、最もマッチング距離の近い単語標準パターンの番号 k と両パターン間の距離 WD_k とを制御部7に送出する。

【0028】こうして、上記登録単語に係る入力音声の認識が終了すると、制御部7によって動作モードが上記話者認識モードに切り替えられる。

【0029】〔話者認識モード〕そうすると、上記話者認識部8においては、上述のようにバッファに格納された所望する電話番号に対応付けられている登録単語あるいは消去/修正の対象となる登録単語のデジタル音声信号に基づいて上記マッチングを行い、最もマッチング距離の近い話者標準パターンの番号 k と両パターン間の距離 SP_k とを制御部7に送出する。

【0030】以下、上記制御部7は、単語認識部1から入力された上記番号 k に基づく最もマッチング距離の近い特徴パターンと単語標準パターンとの距離 WD_k の値(すなわち、距離 WD の最小値)がある閾値より大きい場合には、現在入力された音声の認識に失敗したと判断する。そして、話者認識部8から入力された上記番号 k に基づく最も近いマッチング距離 SP_k を呈する話者標準パターンに対応付けられている単語標準パターンを、上記内部メモリに格納されている上記リンク情報を参照して求め、この求められた単語標準パターンを上記マイクロホン11から音声入力された登録単語の正しい認識結果であると確定する。つまり、上記単語認識部1による登録単語認識に失敗した場合には、話者認識部8によって認識された話者の情報を用いて登録単語を確定するのである。このように、本実施例においては、制御部7で上記認識結果確定部を構成するのである。

【0031】図2は、上記制御部7の制御の下に実施される上記単語認識部1と話者認識部8によるマッチング処理動作および単語標準パターン確定処理動作のフローチャートである。以下、図2に従って上記マッチング処理動作および単語標準パターン確定処理動作について詳細に説明する。動作モードが上記単語認識モードに設定されて、マッチング処理動作および単語標準パターン確定処理動作がスタートする。ステップS1で、単語標準パターン番号 i (最大値は“1”)に初期値“1”がセットされ、変数 k に初期値“1”がセットされる。また、変数 j には十分大きな定数がセットされる。

【0032】ステップS2で、上記単語認識部1によって、得られた入力音声の特徴パターン x と“ i ”番目の単語標準パターンのマッチング距離 $WD(i, x)$ が算出される。ステップS3で、上記算出されたマッチング距離 $WD(i, x)$ の値が変数 j の値より小さいか否かが判別される。その結果小さい場合にはステップS4に進み、そ

うでなければステップS4をスキップする。ステップS4で、上記変数 j にマッチング距離 $WD(i, x)$ の値がセットされる。また、変数 k の値に単語標準パターン番号 i の値がセットされる。

【0033】ステップS5で、単語標準パターン番号 i の内容が最大値“1”より小さいか否かが判別される。その結果“1”より小さければステップS6に進み、そうでなければステップS7に進む。ステップS6で、単語標準パターン番号 i の内容がインクリメントされてステップS2にリターンし、次の単語標準パターンとのマッチングが実施される。ステップS7で、変数 j の内容(すなわち、マッチング距離 $WD(i, x)$ の最小値:上記 WD_k に相当)が閾値“TWD”より小さいか否かが判別される。その結果“TWD”より小さければステップS14に進み、そうでなければステップS8に進む。

【0034】こうして、上記単語認識部1による登録単語認識に失敗すると、動作モードが上記話者認識モードに切り替えられて次のステップに進む。ステップS8で、話者標準パターン番号 i に初期値“1”がセットされ、変数 k に初期値“1”がセットされる。また、変数 j には十分大きな定数がセットされる。ステップS9で、上記話者認識部8によって、上記バッファに格納されたデジタル音声信号に基づいて得られた入力音声の話者特徴パターン y と“ i ”番目の話者標準パターンとのマッチング距離 $SP(i, y)$ が算出される。ステップS10で、上記算出されたマッチング距離 $SP(i, y)$ の値が変数 j の値より小さいか否かが判別される。その結果小さい場合にはステップS11に進み、そうでなければステップS11をスキップする。ステップS11で、上記変数 j にマッチング距離 $SP(i, y)$ の値がセットされ、変数 k の値に単語標準パターン番号 i の値がセットされる。

【0035】ステップS12で、話者標準パターン番号 i の内容が最大値“1”より小さいか否かが判別される。その結果“1”より小さければステップS13に進み、そうでなければステップS14に進む。ステップS13で、話者標準パターン番号 i の内容がインクリメントされてステップS9にリターンし、次の話者標準パターンとのマッチングが実施される。

【0036】ステップS14で、変数 k の値(すなわち、最小のマッチング距離 $SP(i, y)$ (上記 SP_k に相当)を呈する話者特徴パターンの番号)に対応付けられた単語標準パターンが目的とする単語標準パターンであると確定されて、上記マッチング処理動作および単語標準パターン確定処理動作を終了する。このようにして、所望する電話番号に対応付けられている登録単語あるいは消去/修正の対象となる登録単語の単語標準パターンが確定されて選出されるのである。

【0037】つまり、本実施例では、上記マイクロホン11から入力された音声信号に基づいて得られた特徴パターンと単語パターンメモリ2に格納されている i 番目

の単語標準パターンとの重み付け距離MDを

$$MD = A \times WD(i, x) + B \times SP(i, y)$$

但し、 $WD(i, x)$ の最小値 $<TWD$ ならば、 $A = 1$ 、 $B = 0$

$WD(i, x)$ の最小値 $\geq TWD$ ならば、 $A = 0$ 、 $B = 1$ とすると、この重み付け距離MDの値を最小にするi番目の単語標準パターンを選出するのである。

【0038】以後は、上記マイクロホン11から入力された音声が入信時における上記所望する電話番号に対応付けられた登録単語を表す音声である場合には、制御部7は上記確定された単語標準パターンに関連付けられている電話番号を上記内部メモリの内容を参照して得る。そして、得られた電話番号を表す信号を上記回線接続部13に送出する。

【0039】一方、登録単語更新時において入力された消去/修正の対象となる登録単語を表す音声である場合には、制御部7は上記内部メモリに格納されたリンク情報を参照して音声分析/合成部3を制御し、上記確定された単語標準パターンにリンクされたエコーバック音用パラメータを読み出して音声信号を生成する。こうして、上記確定された単語標準パターンのエコーバック音をレシーバ12から出力して、ユーザに対して上記確定された単語標準パターンが目的とする登録単語の単語標準パターンであるかの確認を促すのである。そして、ユーザによって上記レシーバ12からのエコーバック音が目的とする単語標準パターンの発声であると確認されて、上記キー入力部10から消去/修正の指示がキー入力されると、その指示に従って制御部7によって上記所望する単語標準パターンが消去あるいは修正される。

【0040】また、ユーザによって上記レシーバ12からのエコーバック音が目的とする単語標準パターンの発声でないと確認されて、上記キー入力部10からキャンセルの指示がキー入力された場合には、例えば次のような処理を実施する。すなわち、上記キー入力部10からの指示に従って、制御部7によって上記確定された単語標準パターンおよびこれに対応付けられている話者標準パターンをマスクして図2に示すフローチャートを再スタートするのである。こうすることによって、話者認識部8による第2位の認識結果に対応付けられた単語標準パターンが目的とする登録単語の単語標準パターンとして確定される。以下、これを繰り返すことによって目的とする単語標準パターンが検索できるのである。

【0041】その際におけるエコーバック音による検索は、話者認識部8による高順位の認識結果に対応付けられた単語標準パターンから優先的に検索されることになり、エコーバック音による検索範囲を自動的に発声ユーザの単語標準パターンに絞り込むのと同じことになる。したがって、本実施例においては、ユーザ番号指定を実施する必要がないのである。その結果、複数ユーザの単語標準パターンを同じ領域に混在させて登録でき、単語

パターンメモリ2を有効に使用できる。

【0042】次に、上述の動作を具体例を上げて説明する。当該音声ダイヤラ付きの電話機を主人、奥さん、子供の3人で使用し、そのうち奥さんは主人の勤務先の電話番号の登録単語“お父さん”を4番目の登録単語として登録し、子供の学校の電話番号の登録単語“太郎”を5番目の登録単語として登録している。また、主人と子供は夫々3個の登録単語を登録しているものとする。ここで、奥さんは主人の勤務先の電話番号の登録単語を変更する際に、登録単語が“お父さんの会社”であるか“会社”であるか分からなくなった場合を例に上げる。

【0043】(1)奥さんが上記マイクロホン11に向かって「お父さんの会社」と発声した場合
発声単語「お父さんの会社」の単語標準パターンxは4番目の登録単語“お父さん”の単語標準パターンに近い
ため、単語認識部1による各マッチング距離WDと上記閾値“TWD”との大小関係は、
 $TWD > WD(4, x) < \dots < WD(i, x) < \dots < WD(5, x) < \dots$

であるとする。そうすると、図2におけるステップS7、S14で、目的とする登録単語は“お父さん”であると確定される。そして、奥さんは、上記レシーバ12から出力される合成音声「お父さん」によって、変更したい登録単語“お父さん”が検索されたと判断する。そして、キー入力部10から登録単語更新の指示を行う。

【0044】(2)奥さんが上記マイクロホン11に向かって「会社」と発声した場合
発声単語「会社」の単語標準パターンxは両登録単語“お父さん”、“太郎”の単語標準パターンからの距離が遠いため、単語認識部1による各マッチング距離WDと上記閾値“TWD”との大小関係は、
 $TWD < \dots < WD(5, x) < \dots < WD(i, x) < \dots < WD(4, x) < \dots$

であるとする。そうすると、図2におけるステップS7で、目的とする登録単語の単語標準パターンの検索に失敗する。

【0045】さらに、話者認識部8による各マッチング距離SPの大小関係は、

$$SP(5, x) < SP(4, x) < \dots < SP(i, x) < \dots$$

であるとする。そうすると、図2におけるステップS14で、目的とする登録単語は“太郎”であると確定される。そして、奥さんは、上記レシーバ12から出力される合成音声「太郎」によって、変更したい登録単語“お父さん”はまだ検索されていないと判断する。そこで、キー入力部10から登録単語検索の続行を指示する。

【0046】そうすると、5番目の登録単語“太郎”に係る話者標準パターンをマスクして図2のルーチンが再度実施される。その際における話者認識部8による各マッチング距離SPの大小関係は、
 $SP(4, x) < \dots < SP(i, x) < \dots$

となるから、図2におけるステップS14で、目的とする登録単語は“お父さん”であると確定される。そして、奥さんは、上記レシーバ12から出力される合成音声「お父さん」によって、変更したい登録単語“お父さん”が検索されたと判断する。そして、キー入力部10から登録単語更新の指示を行う。

【0047】このように、本具体例によれば、エコーバック音による2回の検索動作で変更したい登録単語“お父さん”が検索される。ところが、従来のユーザ番号指定によらないエコーバック音による検索の場合には、登録単語“お父さん”は4番目の登録単語であるから、4回の検索動作で検索されるのである。

【0048】(3)奥さんが主人の会社の電話番号の登録単語を完全に忘れたために、上記マイクロホン11に向かって「何を入れたのか忘れたので教えて下さい」と発声した場合

この場合、発声単語「何を入れたのか忘れたので教えて下さい」の話者標準パターン y は両登録単語“お父さん”、“太郎”の話者標準パターンから極端に遠い距離にはあるが、話者としては奥さんが一番近いので、話者認識部8による各マッチング距離 SP の大小関係は(2)の場合と同様に奥さんの話者標準パターンとの距離が最も小さくなる。

【0049】したがって、(3)の場合にも(2)の場合と同じ経過をたどって変更したい登録単語“お父さん”が検索されるのである。すなわち、最大2回のエコーバック音による検索動作で目的とする登録単語“お父さん”が検索できる。このように、本実施例によれば、完全に登録単語を忘れても少ない検索回数で短時間に目的とする登録単語を検索できるのである。

【0050】このように、本実施例における音声ダイヤラは、単語認識モード時には、単語をマイクロホン11に向かって発声すると、単語認識部1は、入力された音声信号に基づく特徴パターン x と単語パターンメモリ2に格納された単語標準パターンとのマッチングを実施し、最小マッチング距離 $WD(k, x)$ を呈する単語標準パターンの番号 k とそのマッチング距離 $WD(k, x)$ を制御部7に送出する。また、話者認識モード時には、話者認識部8は、上記単語認識モード時にマイクロホン11に入力された上記単語の音声信号に基づく話者特徴パターン y と話者パターンメモリ9に格納された話者標準パターンとのマッチングを実施し、最小マッチング距離 $SP(k, y)$ を呈する話者標準パターンの番号 k とそのマッチング距離 $SP(k, y)$ を制御部7に送出する。

【0051】そして、上記制御部7は、マッチング距離 $WD(k, x)$ の値が所定値“ TWD ”より小さい場合には、単語認識部1から入力された番号 k の単語標準パターンを入力音声の正しい認識結果であると確定する。一方、マッチング距離 $WD(k, x)$ の値が上記所定値“ TWD ”以上である場合には、話者認識部8から入力された番号 k

の話者標準パターンに関連付けられた単語標準パターンを入力音声の正しい認識結果であると確定する。

【0052】こうすることによって、当該音声ダイヤラが設けられた電話機を複数のユーザが使用する場合であっても、音声認識部1によって目的とする登録単語の単語標準パターンを検索できない場合には、自動的にエコーバック音による検索範囲を発声ユーザの単語標準パターンに絞り込むことができる。したがって、ユーザ番号指定によらずとも、自動的に発声ユーザの単語標準パターンの中から短時間に目的とする登録単語の単語標準パターンを検索することができるのである。

【0053】また、上述のように、エコーバック音による検索範囲を自動的に発声ユーザの単語標準パターンに絞り込むことができるので、目的とする登録単語を忘れた場合であっても、マイクロホン11に向かって何らかの発声を行えば発声ユーザの登録単語のエコーバック音を得ることができる。したがって、登録単語を忘れても短時間に目的とする登録単語を検索できるのである。

【0054】上記実施例においては、図2のフローチャートに基づく一回の単語標準パターン検索において、最もマッチング距離の小さい単語標準パターンのみを出力するようにしている。しかしながら、この発明はこれに限定されるものではなく、マッチング距離の小さい順に複数の単語標準パターンを正しい認識結果であるとして確定して、複数の認識候補を出力するようにしてもよい。

【0055】また、上記実施例においては、上述のようにして検索された単語標準パターンを音声分析/合成部3で生成された合成音によるエコーバック音によって確認するようにしている。しかしながら、この発明はこれに限定されるものではなく、検索された単語標準パターンの登録単語を表示ディスプレイに表示するようにしてもよい。特に、目的とする登録単語を全く忘れてしまった場合には、上記マッチング距離 $SP(i, x)$ の小さい発声ユーザの話者標準パターンに係る複数の登録単語を認識候補として一度に画面表示することによって、より迅速に目的とする登録単語を検索できる。

【0056】また、この発明における単語標準パターン検索のアルゴリズムは図2のフローチャートに限定されるものではない。

【0057】

【発明の効果】以上より明らかなように、第1の発明の音声ダイヤラは、認識結果確定部によって、上記音声認識部による認識結果に係る最小マッチング距離が所定値以上である場合には上記話者認識部による認識結果に関連付けられた単語標準パターンを入力音声の正しい認識結果であると確定し、確定した単語標準パターンの登録単語を認識結果出力部によって出力するので、上記音声認識部による単語標準パターンの検索に失敗した場合には自動的に検索範囲を発声ユーザの単語標準パターンに

絞り込むことができる。したがって、この発明の音声ダイヤラによれば、ユーザ番号指定によらずとも自動的に検索範囲を発声ユーザの単語標準パターンに限定して目的とする単語標準パターンを短時間に検索できる。

【0058】さらに、上述のように、上記音声認識部による単語標準パターンの検索に失敗した場合には自動的に検索範囲を発声ユーザの単語標準パターンに絞り込むことができるので、目的とする登録単語を忘れた場合でも、発音の似た単語を発声することによって目的とする登録単語の単語標準パターンを短時間に検索できる。

【0059】また、第2の発明の音声ダイヤラは、上記認識結果確定部は、上記音声認識部からの複数の単語標準パターンによる認識結果と上記話者認識部からの複数の話者標準パターンによる認識結果に基づいて入力音声の正しい認識結果を確定し、上記認識結果出力部は上記確定された複数の単語標準パターンの登録単語を認識候補として出力するので、第1の発明の効果に加えて、目的とする登録単語を全く忘れた場合でも、何らかの単語

を発声すれば発声ユーザの単語標準パターンによる認識候補が上記認識出力部によって出力される。したがって、この発明によれば、登録単語を完全に忘れてしまった場合であっても目的とする登録単語の単語標準パターンを短時間に検索できる。

【図面の簡単な説明】

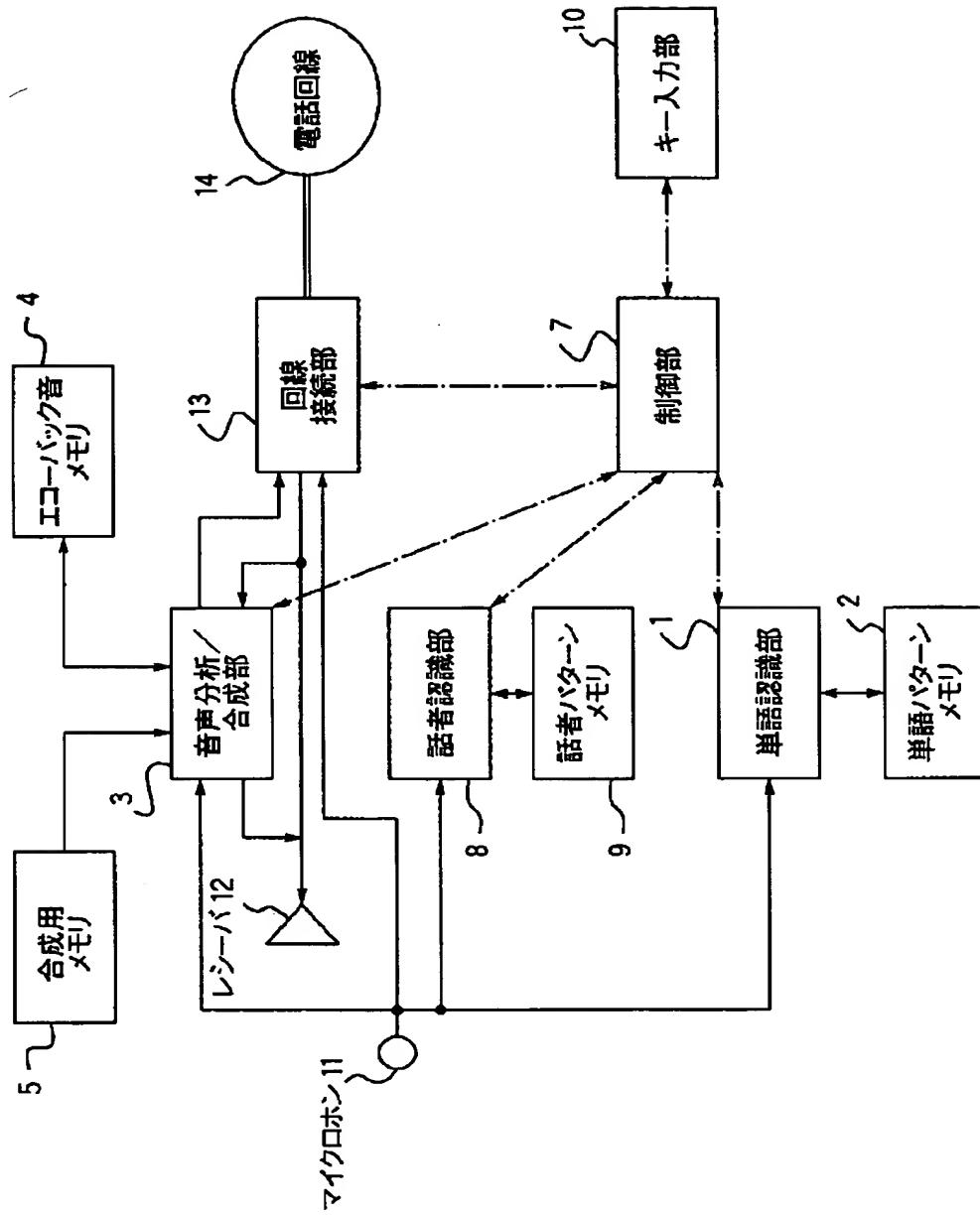
【図1】この発明の音声ダイヤラにおける一実施例を示すブロック図である。

【図2】図1における制御部の制御の下に実施されるマッチング処理動作および単語標準パターン確定処理動作のフローチャートである。

【符号の説明】

1…単語認識部、2…単語パターンメモリ、3…音声分析/合成部、4…エコーバック音メモリ、7…制御部、8…話者認識部、9…話者パターンメモリ、10…キー入力部、13…回線接続部、14…電話回線。

【図1】



【図2】

